exposing the nucleotide sequence to an effective amount of the composition of any of claims 2 to 7.

Please amend claim 15, without prejudice, as follows:

15. (Amended) The process of claim 14 wherein the sequence 5'-$(GNN)_n$-3' (SEQ ID NO:123) is located in the transcribed region of the nucleotide sequence.

Please amend claim 16, without prejudice, as follows:

16. (Amended) The process of claim 14 wherein the sequence 5'-$(GNN)_n$-3' (SEQ ID NO:123) is located in a promotor region of the nucleotide sequence.

Please amend claim 17, without prejudice, as follows:

17. (Amended) The process of claim 14 wherein the sequence 5'-$(GNN)_n$-3' (SEQ ID NO:123) is located within an expressed sequence tag.

REMARKS

The amendments detailed above provide unique sequence identifiers for each sequence disclosed in the specification, claims, and corresponding sequence listing as required under 37 C.F.R. §§ 1.1821 - 1.825. The amendments to the specification are supported by the sequences disclosed in the specification itself. The amendments to the claims are supported by the sequences disclosed in the claims themselves. Applicants believe that no new matter is added by way of the present response.

Applicants respectfully request entry of the amendments and sequence listing provided herein and in supporting documents, including the CRF.

Attached hereto is Appendix I which is a marked-up version of the changes made to the specification and claims by the current amendment.

Applicants believe that the present response is a bona fide attempt to advance the application, and is a complete reply to the Notice to Comply with Requirements For Patent Applications Containing Nucleotide sequence and/or Amino Acid Sequence Disclosures mailed March 04, 2003. Therefore, if consideration of some matter or compliance with some requirement has been inadvertently omitted, Applicants respectfully request that the Examiner provide a new time period for reply under 37 C.F.R. § 1.134 to supply the omission.

The Examiner is requested to contact the representative for the Applicants, to discuss any questions or for clarification. If there are any further fees associated with this response, the Director is authorized to charge our Deposit Account No. 19-0962.

Respectfully submitted,

April 02, 2003
Date

Michael J. McCarthy, Reg. No. 46,910

THE SCRIPPS RESEARCH INSTITUTE
10550 North Torrey Pines Road
Mail Drop TPC-8
La Jolla, California 92037
(858) 784-2937

APPENDIX I

VERSION WITH MARKINGS TO SHOW CHANGES MADE

In the Specification:

The paragraph beginning at page 5, line 19, has been amended as follows:

In a still further aspect, the present invention provides a process of regulating the function of a nucleotide sequence that contains the sequence 5'-(GNN)n-3' (SEQ ID NO:123), where n is an integer from 1 to 6, the process comprising exposing the nucleotide sequence to an effective amount of a composition of this invention operatively linked to one or more transcription modulating factors. The 5'-(GNN)n-3' sequence can be found in the transcribed region or promotor region of the nucleotide or within an expressed sequence tag.

The paragraph beginning at page 10, line 20, has been amended as follows:

From the combined selection and mutagenesis data it emerged that specific recognition of many nucleotides could be best accomplished using motifs, rather than a single amino acid. For example, the best specification of a 3' guanine was achieved using the combination of Arg-1, Ser1, and Asp2 (the RSD motif). By using Val5 and Arg6 to specify a 5' guanine, recognition of subsites GGG, GAG, GTG, and GCG could be accomplished using a common helix structure (SRSD-X-LVR) (SEQ ID NO:124) differing only in the position 3 residue (Lys3 for GGG, Asn3 for GAG, Glu3 for GTG, and Asp3 for GCG). Similarly, 3' thymine was specified using Thr-1, Ser1, and Gly2 in the final clones(the TSG motif).

Further, a 3' cytosine could be specified using Asp-1, Pro1, and Gly2 (the DPG motif) except when the subsite was GCC; Pro1 was not tolerated by this subsite. Specification of a 3' adenine was with Gln-1, Ser1, Ser2 in two clones (QSS motif). Residues of positions 1 and 2 of the motifs were studied for each of the 3' bases and found to provide optimal specificity for a given 3' base as described here.

The paragraph beginning at page 13, line 24, has been amended as follows:

The data show that all possible GNN triplet sequences can be recognized with exquisite specificity by zinc finger domains. Optimized zinc finger domains can discriminate single base differences by greater than 100-fold loss in affinity. While many of the amino acids found in the optimized proteins at the key contact positions -1,3, and 6 are those that are consistent with a simple code of recognition, it has been discovered that optimal specific recognition is sensitive to the context in which these residues are presented. Residues at positions 1,2, and 5 have been found to be critical for specific recognition. Further the data demonstrates for the first time that sequence motifs at positions -1,1, and 2 rather than the simple identity of the position 1 residue are required for highly specific recognition of the 3' base. These residues likely provide the proper stereochemical context for interactions of the helix both in terms of recognition of specific bases and in the exclusion of other bases, the net result being highly specific interactions. Broad utility of these domains would be realized if they were modular in both their interactions with DNA and other zinc finger

domains. This could be achieved by working within the likely limitations imposed by target site overlap, namely that sequences of the 5'-(GNN)$_x$-3' type should be targeted. Ready recombination of the disclosed domains then allows for the creation of polydactyl proteins of defined specificity precluding the need to develop phage display libraries in their generation. These polydactyl proteins have been used to activate and repress transcription driven by the human *erbB-2* promoter in living cells. The family of zinc finger domains described herein is likely sufficient for the construction of 16^6 or 17 million novel proteins that bind the 5'-(GNN)$_6$-3' <u>(SEQ ID NO:125)</u> family of DNA sequences.

The paragraph beginning at page 20, line 13, has been amended as follows:

The *erbB-2* promoter therefore represents an interesting test case for the development of artificial transcriptional regulators. This promoter has been characterized in detail and has been shown to be relatively complex, containing both a TATA-dependent and a TATA-independent transcriptional initiation site (Ishii, S., Imamoto, F., Yamanashi, Y., Toyoshima, K. & Yamamoto, T. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 4374-4378). Whereas early studies showed that polydactyl proteins could act as transcriptional regulators that specifically activate or repress transcription, these proteins bound upstream of an artificial promoter to six tandem repeats of the proteins binding site (Liu, Q., Segal, D. J., Ghiara, J. B. & Barbas III, C. F. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 5525-5530). Furthermore, this study utilized polydactyl proteins that were not modified in their

binding specificity. Herein, we tested the efficacy of polydactyl proteins assembled from predefined building blocks to bind a single site in the native *erbB-2* promoter. Described above is the generation and characterization of a family of zinc finger domains that bind each of the 16 5'-GNN-3' DNA triplets. One reason we focused on the production of this family of recognition domains is that promoter regions of most organisms are relatively GC rich in their base content. Thus, if proteins recognizing 5'-$(GNN)_x$-3' sites could be readily assembled from this set of defined zinc finger domains, many genes could be rapidly and specifically targeted for regulation. A protein containing six zinc finger domains and recognizing 18 bp of DNA should be sufficient to define a single address within all known genomes. Examination of the *erbB-2* promoter region revealed two 5'-$(GNN)_6$-3'sites <u>(SEQ ID NO:125)</u> and one 5'-$(GNN)_4$-3' <u>(SEQ ID NO:126)</u> site. One of these sites, identified here as e2c, falls within the 5'-untranslated region of the *erbB-2* gene and was chosen as the target site for the generation of a gene-specific transcriptional switch. A BLAST sequence similarity search of the GenBank data base confirmed that this sequence is unique to *erbB-2*. The position of the e2c target sequence, downstream and in the vicinity of the two major transcription initiation sites, allowed for the examination of repression through inhibition of either transcription initiation or elongation. An interesting feature of the e2c target site is that it is found within a short stretch of sequence that is conserved between human, rat, and mouse *erbB-2* genes (White, M. R.-A. & Hung, M.-C. (1992) *Oncogene* **7**, 677-683). Thus, targeting of this site would allow for the study of this strategy in animal models prior to its application to human disease.

The paragraph beginning at page 21, line 20, is amended as follows:

The general utility of two different strategies for generating three-finger proteins recognizing 9 bp of DNA sequence was investigated. Each strategy was based on the modular nature of the zinc finger domain, and takes advantage of a family of zinc finger domains recognizing triplets of the 5′-GNN-3′. Two three-finger proteins recognizing halfsites (HS) 1 and 2 of the 5′-(GNN)[illegible]-3′ (SEQ ID NO:125) *erbB-2* target site e2c were generated in the first strategy by fusing the pre-defined finger 2 (F2) domain variants together using a PCR assembly strategy. To examine the generality of this approach, three additional three-finger proteins recognizing sequences of the 5′-(GNN)[illegible]-3′ type, were prepared using the same approach. Purified zinc finger proteins were prepared as fusions with the maltose binding protein (MBP). ELISA analysis revealed that serially connected F2 proteins were able to act in concert to specifically recognize the desired 9-bp DNA target sequences. Each of the 5 proteins shown was able to discriminate between target and non-target 5′-(GNN)[illegible]-3′ sequence.

The paragraph beginning at page 22, line 29, has been amended as follows:

As discussed above, the recognition of 9 bp of DNA sequence is not sufficient to specify a unique site within a complex genome. In contrast, a six-finger protein recognizing 18 bp of contiguous DNA sequence could define a single site in the human genome, thus fulfilling an important prerequisite for the

generation of a gene-specific transcriptional switch. Six-finger proteins binding the erbB-2 target sequence e2c were generated from three-finger constructs by simple restriction enzyme digestion and cloning with F2, Zif268, and Sp1C framework template DNAs. ELISA analysis of purified MBP fusion proteins showed that each of the six-finger proteins was able to recognize the specific target sequence, with little cross reactivity to non-target 5′-$(GNN)_6$-3′ <u>(SEQ ID NO:125)</u> sites or a tandem repeat of the Zif268 target site.

<u>In the Claims:</u>

Claim 7 has been amended as follows:

7. (Amended) The composition of any of claims 2 to 6 that binds to a nucleotide that contains the sequence 5'-$(GNN)_n$-3' <u>(SEQ ID NO:127)</u>, wherein each N is A, C, G, or T with the proviso that all N′s cannot be C and where n is 2 to 6.

Claim 14 has been amended as follows:

14. (Amended) A process of regulating a nucleotide sequence that contains the sequence 5'-$(GNN)_n$-3' <u>(SEQ ID NO:123)</u>, where n is an integer from 1 to 6, the process comprising exposing the nucleotide sequence to an effective amount of the composition of any of claims 2 to 7.

Claim 15 has been amended as follows:

15. (Amended) The process of claim 14 wherein the sequence 5'-

$(GNN)_n$-3' (SEQ ID NO:123) is located in the transcribed region of the nucleotide sequence.

Claim 16 has been amended as follows:

16. (Amended) The process of claim 14 wherein the sequence 5'-$(GNN)_n$-3' (SEQ ID NO:123) is located in a promotor region of the nucleotide sequence.

Claim 17 has been amended as follows:

17. (Amended) The process of claim 14 wherein the sequence 5'-$(GNN)_n$-3' (SEQ ID NO:123) is located within an expressed sequence tag.